DOCUMENT RESUME

ED 141 381 TM 006 207

AUTHOR Hambleton, Ronald K.; Murray, John
TITLE A Comparative Study of Faculty and Student Attitudes Toward a Variety of College Grading Purposes and Practices.
PUB DATE [Apr 77]
NOTE 50p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (New York, New York, April 5-7, 1977) ; Tables may be marginally legible due to small type

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *College Faculty; College Majors; *College Students; Educational Research; Grades (Scholastic); *Grading; Guidelines; Higher Education; Learning; Pass Fail Grading; *Student Attitudes; Student Characteristics; *Teacher Attitudes; Teacher Background

ABSTRACT

Of the many issues facing higher education, perhaps none has been more frequently and hotly debated by college administrators, instructors, and students, than the issue of college grading purposes and practices. Regretably, much of the research to date has been poorly done and hence has led to few changes. The purpose of this research was to overcome many of the past deficiencies and provide a comprehensive study of faculty and students' views concerning the uses of grading in several instructional settings, and the appropriateness of a variety of commonly used grading systems for accomplishing intended uses of grades. Specifically, faculty and students' views on: (1) The importance of twelve possible uses of grades in different instructional settings (courses in students' major area of concentration versus non-major areas of concentration); (2) The acceptability of each of the five common grading systems for accomplishing twelve possible uses of grades; and (3) The effects of each of five common grading systems on a variety of course outcomes (for example, maintaining academic standards and maximizing amount of learning). Overall, the results strongly support the belief that faculty and students are in favor of a criterion-referenced grading system. While the results cannot be generalized to other institutions, several innovations in the research design should provide guidelines for researchers to enable them to conduct better studies on grading in their own institutions. (Author/MV)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED141381

A Comparative Study of Faculty and Student Attitudes Toward A Variety of College Grading Purposes and Practices

Ronald K. Hambleton
University of Massachusetts, Amherst

and

John Murray
Boston State College

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Abstract

Of the many issues facing Higher Education, perhaps none has been more frequently and hotly debated by college administrators, instructors, and students, than the issue of college grading purposes and practices. Regretably, much of the research to date has been poorly done and hence has lead to few changes. The purpose of this research was to overcome many of the past deficiencies and provide a comprehensive study of faculty and students' views concerning the uses of grading in several instructional settings, and the appropriateness of a variety of commonly used grading systems for accomplishing intended uses of grades. Overall, the results strongly support the belief that faculty and students are in favor of a criterion-referenced grading system. While the results cannot be generalized to other institutions, several innovations in the research design should provide guidelines for researchers to enable them to conduct better studies on grading in their own institutions.

TM006 207

A Comparative Study of Faculty and Student Attitudes Toward A Variety of College Grading Purposes and Practices[1,2]

Ronald K. Hambleton
University of Massachusetts, Amherst

and

John Murray
Boston State College

The issue of college grading practices has been frequently debated among college administrators, instructors, and students. That the issue is an important one is clear when one recognizes that grades directly affect the career choices of many students (Baucom, 1974; Stevens, 1973; Warren, 1971). In addition, there is substantial evidence to suggest that college grading practices affect student attitudes toward learning, the amount of learning, course selections, and the amount of time spent in study (Warren, 1971).

Unfortunately though, because of the complexity of the college grading practices issue, and the inexperience of most college administrators and instructors in the areas of test development and educational measurement, considerable confusion exists about proper grading practices. Contributing to the confusion is a plethora of research studies that have only limited generalizability and/or have been rather narrowly focused on the topics of symbolic representation of grades, notational systems, and the use of grades for prediction. Also, on the most important question, that concerning the purposes or uses of grades, there has been only limited research. In recent papers, both Hambleton and Rovinelli (in press) and Williams and Miller (1973) have stressed the importance of addressing the uses of grading as a first step in a general plan for comprehensive study of college grading purposes and practices.

---

[1]A paper presented at the annual meeting of NCME, New York, 1977.

[2]Laboratory of Psychometric and Evaluative Research Report No. 48. Amherst, Mass.: School of Education, University of Massachusetts, 1977.

The research described in this paper was designed to compare faculty and students' views on the following questions:

1. What is the importance of twelve possible uses of grades in different instructional settings (courses in students' major area of concentration versus non-major areas of concentration)?

2. What is the acceptability of each of five common grading systems for accomplishing twelve possible uses of grades?

3. What are the effects of each of five common grading systems on a variety of course outcomes (for example, maintaining academic standards and maximizing amount of learning)?

This research project was designed to respond to several shortcomings of many earlier studies. First, researchers have seldom considered the dimension of "instructional setting" in their work. It is possible that instructors' and students' views of the uses of grades and appropriate grading systems will vary from one instructional setting to another. Second, too often researchers have compared one grading system with another without regard for the intended uses of grades in a particular instructional setting. It is important to study the merits of different grading systems against some very different (and common) uses of grades, and many important course outcomes. Corresponding to each intended use or purpose of grading are appropriate methods of test design and grading. The value of different grading systems is situation-specific--change the intended use or purpose of grades and the merits of different grading systems will also shift. Finally, it is important to consider both faculty and students' views. The ideal grading system would be one that was responsive to the intended purposes or uses of grades in a particular course and was seen as appropriate by both faculty and students.

Method

Instrumentation

Data for the study were collected from the administration of two questionnaires (one designed for faculty members and the other one designed for students). The two questionnaires were essentially identical, differing only in the background questions that were asked of faculty members and students. It took about 40 minutes to complete. In an attempt to obtain high return rates the purposes of the study were carefully explained to faculty and students. Also, the potential policy implications of the results were stressed.

The questionnaires were divided into four parts.[1] The four parts will be discussed next. (Copies of the questionnaires may be obtained by writing the senior author and enclosing a stamped and self-addressed envelope.)

Part A—Background

This section differed for faculty and students. Background questions for the faculty members included questions about their age, rank, sex, years of teaching experience, types of classes taught, and primary teaching area. Students were asked questions concerning their sex, age, student status, grades in college, SAT scores, experience with P/F grading, major area of concentration, and future plans.

Part B—Uses of Grades in Different Instructional Settings

In this section, faculty and students were asked to complete the following task:

> Many uses of grades have been suggested and some uses are more important than others in some types of courses and with certain

[1]Actually there were two other parts of the questionnaire: One that was designed to permit us to study the factors affecting grades in six different instructional settings. The other part consisted of a series of open-ended questions about grading. This part provided faculty and students with an opportunity to indicate additional comments. The data derived from these two partsare reported in Murray (1977).

types of students. For example, an instructor may view the uses of grades differently for students taking courses in their major area of concentration than for students taking an elective course in an area different from their major area of concentration.

Below are listed 12 uses of grades and two types of instructional settings: Students taking courses in their major area of concentration, and non-major areas of concentration, respectively. Your task is to indicate how important you feel each use of grades is in each instructional setting. There are four possible ratings:

1=very important, 2=important, 3=somewhat important, 4=not important.

Under investigation are the following twelve uses of grades:

1. Motivate students to do good work in the course.
2. Provide instructors with information about student progress.
3. Improve a student's ability to critically assess his or her own work.
4. Provide instructors with information about their teaching methods.
5. Provide instructors with information about their curriculum.
6. Assess teacher competence.
7. Rank students on their performance in the course.
8. Inform others (e.g., advisors, future teachers, employers, registrar) about student performance.
9. Compare student performance against absolute standards of performance.
10. Summarize multiple evaluations of a student during a course into a single digit or letter.
11. Provide students with feedback on their course performance.
12. Provide a method for maintaining academic standards.

The twelve uses were generated from a review of many past studies on grading practices (Ericksen & Bluestone, 1971; Hunt, 1972; Karlins, Kaplan, & Stuart, 1969; Scriven, 1976; Warren, 1971).

Part C-Relative Merits of Different Methods of Grading

Five grading systems were selected for investigation: Pass/Fail, Pass/No Record, Honors/Pass/Fail, Norm-referenced Grading, Criterion-referenced Grading. The last (criterion-referenced grading) is perhaps the least known among the five, but it is currently attracting considerable attention from instructors at all levels of education. Test construction and test score interpretations of criterion-referenced tests have been

discussed by many researchers (see, for example, Hambleton & Novick, 1973; Millman, 1974).

The five grading systems were introduced to faculty and students with the following definitions:

Pass/Fail

In this grading system, the student either passes or fails a course. For a passing grade, the student receives credit toward graduation. For a failing grade, the student receives no credit toward graduation and a failing grade on his/her transcript.

Pass/No Record

This grading system is similar to Pass/Fail except that failing grades are not recorded on a student's transcript. When a student fails a course, there will be no record that the student had even been enrolled in the course.

Honors/Pass/Fail

The grading system is similar to Pass/Fail grading except that among passing students a distinction is made between "superior" and "passing" performance.

Norm-referenced Grading (Grading on the Curve)

In this grading system, letters A, B, C, D, and F (or some variation) are used to designate the relative performance of students. Letter grades are assigned to reflect student performance with respect to the performance of other students in the course.

Criterion-referenced Grading

In this grading system, letters A, B, C, D, and F (or some variation) are used to designate different levels of course performance. The letter grades are assumed to reflect different levels of performance in some absolute sense. Grades are assigned to students to reflect their level of performance. The important point is that each student is judged on his or her own merits with respect to the standards set by the instructor and not with respect to the performance of other students in the class.

These five grading systems are among the most common in use today. Several other less common grading systems were identified (for example, anecdotal records and Honors/Pass/No Record) but they were excluded to make the rating task in

this, and other parts of the questionnaire, more manageable. As it was, this part of the questionnaire required respondents to make 60 ratings.

The task was for faculty members and students to rate the acceptability (1=highly acceptable, 2=acceptable, 3=minimally acceptable, 4=unacceptable) of each of the five grading systems for accomplishing each of the twelve uses of grades.

Part D-Effects of Different Grading Systems on Various Course Outcomes

The effects of grading on various course outcomes have been a frequent topic of study (for example, Hales, Bain, & Rand, 1973; Hambleton and Rovinelli, 1976; Karlins, et al., 1969; Reiner & Jung, 1972; Stallings & Smock, 1971; Stallings, Wolff & Maehr, 1969). From these and other studies, we identified a list of thirteen course outcomes that have often been discussed as important:

1. Minimize student competition
2. Maximize student enthusiasm
3. Maximize student performance
4. Minimize cheating
5. Maximize outside reading
6. Develop positive self-image
7. Minimize course dropouts
8. Provide data for course evaluation
9. Maintain academic standards
10. Maximize class attendance
11. Maximize expressions of personal opinions
12. Maximize amount of learning
13. Maximize opportunity for further study.

There were three tasks for faculty members and students:

1. Indicate the grading systems which they felt were most likely to help accomplish each course outcome.
2. Indicate the grading systems which they felt were most likely to interfere with the accomplishment of each course outcome.
3. Indicate the course outcomes that were important in their courses.

Sample

All faculty members of the School of Education at the University of Massachusetts were asked to participate in the study. The return rate of questionnaires was 77% (75 of 98). A random sample of faculty members in other parts of the University was also drawn. The return rate of questionnaires from this group was 39% (only 29 of 75). In addition, we selected (at random) ten School of Education and ten University courses in which to administer the student questionnaires. Several of the courses selected were not available to us and so they were replaced by alternate courses. Students in the selected courses were asked to complete the questionnaires at home and return them. The return of questionnaires among students was approximately 50%.

The data were collected in the Spring and Fall of 1976.

Results and Discussion

Introduction

This particular research project was designed to provide data for administration, faculty, and students in the School of Education at the University of Massachusetts, Amherst. The data will be used to assist them in developing new grading systems; grading systems that reflect the educational philosophy of the School. However, for our purposes here, we will discuss the results obtained from faculty members and students in both the School of Education and other parts of the University.

Background

Faculty and student responses to the background questions are summarized in Tables 1 and 2. With respect to the faculty, 74% were males; the majority of the faculty (again, 74%) were in the age range from 31 to 50; 85% had four or more years of college teaching experience; about 50% of the faculty taught both graduate and undergraduate courses, 25% of the faculty taught only graduate students, and the remaining 25% taught only undergraduates; roughly 1/3 of the faculty were in each of three academic rank (Assistant, Associate, Full); and a somewhat surprising 83% of the faculty felt "knowledgeable about methods for assigning grades to students."

With respect to the students, 66% of the students were female; 85% were in the age range 19-27; 23% of the students were freshmen or sophomores, 56% were juniors or seniors, and the remaining 21% were graduate students and special students; and over 60% of the students had had some experience with pass/fail grading. Some additional background statistics are also reported in Table 2.

Importance of 12 Possible Uses of Grades in Different Instructional Settings

Reported in Table 3 are the means and standard deviations of faculty and student ratings of the importance of 12 possible uses of grades in courses in students' major areas of concentration. Surprisingly, these ratings were nearly identical with faculty and student ratings in courses in students' non-major areas of concentration. (Therefore, they are not reported here.) Apparently faculty and students made no distinctions among different uses of grading in students' major and non-major areas of concentrations.

Several statistical methods are available for analyzing the data reported in Table 3. For example, we could have done a multivariate

Table 1

Background Information on School of Education and Non-School of Education Faculty

Questions	Faculty		
	Education (N=75)	Non-Education (N=29)	Total (N=104)
1. What is your sex?			
(1) Male	74.7	71.4	73.8
(2) Female	25.3	28.6	26.2
	χ^2=.1, p=.94		
2. What is your age?			
(1) 21-30	9.3	10.7	9.7
(2) 31-40	40.0	42.9	40.8
(3) 41-50	34.7	28.6	33.0
(4) 51-60	13.3	7.1	11.7
(5) over 60	2.7	10.7	4.9
	χ^2=3.7, p=.45		
3. How many years of experience have you had as a member of a university faculty?			
(1) 0-3 years	20.0	3.4	15.4
(2) 4-6 years	17.3	34.5	22.1
(3) 7-9 years	22.7	6.9	18.3
(4) 10-12 years	13.3	13.8	13.5
(5) over 12 years	26.7	41.4	30.8
	χ^2=10.8, p<.05		
4. What types of classes do you teach?			
(1) mainly graduate level (75% and over)	33.3	3.6	25.2
(2) mainly undergraduate-level (75% and over)	22.7	35.7	26.2
(3) a mixture of graduate-level and undergraduate-level (26-74%)	44.0	60.7	48.5
	χ^2=9.7, p<.01		

Table 1 (cont)

Questions	Faculty		
	Education (N=75)	Non-Education (N=29)	Total (N=104)
5. What is your faculty rank?			
(1) Lecturer	6.8	10.7	7.9
(2) Assistant Professor	34.2	32.1	33.7
(3) Associate Professor	34.2	25.0	31.7
(4) Professor	24.7	32.1	26.7
χ^2=1.4, p=.71			
7. To what extent do you feel knowledgeable about methods for assigning grades to students?			
(1) To a great extent	40.0	29.6	37.3
(2) To a considerable extent	41.3	59.3	46.1
(3) To a limited extent	16.0	11.1	14.7
(4) Not at all	2.7	0.0	2.0
χ^2=3.0, p=.39			
8. Would you be interested in attending workshops to discuss the issue of college grading practices?			
(1) Definitely yes	15.7	7.4	13.4
(2) Yes	28.6	22.2	26.8
(3) No	37.1	40.7	38.1
(4) Unsure	18.6	29.6	21.6
χ^2=2.5, p=.48			

Table 2

Background Information on School of Education and Non-School of Education Students

Questions	Students		
	Education (N=113)	Non-Education (N=336)	Total (N=449)
1. What is your sex?			
(1) male	29.7	35.8	34.3
(2) female	70.3	64.2	65.7
	χ^2=1.1, p=.29		
2. What is your age?			
(1) 18 and under	.9	3.6	2.9
(2) 19-21	46.0	59.7	56.3
(3) 22-24	15.0	19.1	18.1
(4) 25-27	15.9	8.4	10.3
(5) 28-30	7.1	5.4	5.8
(6) over 30	15.0	3.0	6.7
	χ^2=26.5, p<.001		
3. What is your student classification?			
(1) Freshman or sophmore	14.2	25.9	22.9
(2) Junior or senior	46.9	58.6	55.7
(3) First year graduate student	11.5	4.2	6.0
(4) Advanced-level graduate student	21.2	5.1	9.1
(5) Special student	6.2	6.3	6.2
	χ^2=39.0, p<.001		
4. What is your current grade point average?			
(1) less than 2.0	0.9	0.3	0.5
(2) 2.0 - 2.49	7.3	7.2	7.2
(3) 2.5 - 2.99	15.5	24.6	22.3
(4) 3.0 - 3.49	38.2	41.1	40.4
(5) 3.5 - 4.00	27.3	22.8	23.9
(6) don't remember	10.9	3.9	5.6
	χ^2=11.8, p<.05		

Table 2 (cont)

Questions	Students		
	Education (N=113)	Non-Education (N=336)	Total (N=449)
5. What was your SAT verbal score?			
(1) less than 400	3.6	3.0	3.2
(2) 400-500	27.9	28.7	28.5
(3) 501-600	28.8	33.8	32.4
(4) 601-700	14.4	18.1	17.2
(5) over 700	2.7	2.7	2.7
(6) don't remember	22.5	13.9	16.1
	χ^2=5.2, p=.39		
6. What was your SAT quantitative score?			
(1) less than 400	1.8	0.9	1.1
(2) 400-500	17.4	15.2	15.7
(3) 501-600	29.4	30.6	30.3
(4) 601-700	14.7	21.2	19.6
(5) over 700	4.6	11.5	9.8
(6) don't remember	32.1	20.6	23.5
	χ^2=11.4, p<.05		
7. Approximately what percent of your courses have been graded - Pass/Fail?			
(1) about 100%	14.3	1.5	4.7
(2) about 80%	5.4	0.6	1.8
(3) about 60%	9.8	0.3	2.7
(4) about 40%	22.3	2.7	7.6
(5) about 20%	32.1	43.3	40.5
(6) about 0%	16.1	51.6	42.7
	χ^2=138.2, p<.001		
8. How many courses at the University have you taken on a Pass/Fail basis?			
(1) 0	41.1	38.1	38.8
(2) 1-3	47.3	51.2	50.2
(3) 4-6	8.9	9.8	9.6
(4) 7-9	1.8	0.6	0.9
(5) 10 or more	0.9	0.3	0.4
	χ^2=2.5, p=.64		

Table 2 (cont)

Questions	Students		
	Education (N=113)	Non-Education (N=336)	Total (N=449)
9. How many courses have you taken on a Pass/Fail basis in the School of Education?			
(1) 0	12.6	64.1	51.2
(2) 1-3	29.7	29.3	29.4
(3) 4-6	18.0	4.8	8.1
(4) 7-9	14.4	1.2	4.5
(5) 10 or more	25.2	0.3	6.5
	χ^2=173.2, p<.001		
11. Are you planning to go to graduate school?			
(1) definitely	8.9	16.8	14.8
(2) probably	20.5	34.1	30.7
(3) uncertain	25.9	24.0	24.4
(4) probably not	8.0	12.6	11.4
(5) definitely not	3.6	3.0	3.1
(6) I am currently in graduate school	32.1	9.6	15.2
	χ^2=41.2, p<.001		
12. Are you planning to go to professional school?			
(1) definitely	1.8	6.7	5.5
(2) probably	9.8	9.1	9.3
(3) uncertain	23.2	29.6	28.0
(4) probably not	25.0	32.0	30.2
(5) definitely not	19.6	15.9	16.8
(6) I am currently in a professional school	20.5	6.7	10.2
	χ^2=22.6, p<.001		

Table 3

Descriptive Statistical Analysis of Faculty and Student Ratings[1] of the Importance of 12 Possible Uses of Grades in Courses in Students' Major Areas of Concentration

Uses of Grades	Faculty						Students					
	Education (N=75)		Non-Education (N=29)		Total (N=104)		Education (N=113)		Non-Education (N=336)		Total (N=449)	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
1. Motivate students to do good work in the course	2.5	1.1	2.1	1.0	2.4	1.1	2.5	1.2	2.1	1.0	2.2	1.0
2. Provide instructors with information about student progress	2.6	1.2	2.1	1.0	2.4	1.1	2.6	1.1	2.2	0.9	2.3	1.0
3. Improve a student's ability to critically assess his or her own work	2.6	1.1	2.0	1.0	2.4	1.1	2.6	1.1	2.3	1.0	2.4	1.0
4. Provide instructors with information about their teaching methods	3.0	1.2	3.0	1.1	3.0	1.2	2.7	1.1	2.4	1.1	2.5	1.1
5. Provide instructors with information about their curriculum	3.3	0.9	3.3	1.0	3.3	1.0	2.8	1.1	2.6	1.1	2.6	1.1
6. Assess teacher competence	3.2	0.9	3.3	1.0	3.2	0.9	2.8	1.3	2.7	1.1	2.7	1.1
7. Rank students on their performance in the course	3.0	1.1	2.3	1.1	2.8	1.1	2.9	1.0	2.5	1.0	2.6	1.0
8. Inform others (e.g., advisors, future teachers, employers, registrar) about student performance	2.3	1.0	2.0	0.8	2.2	1.0	2.5	1.0	2.4	1.0	2.4	1.0
9. Compare student performance against absolute standards of performance	3.1	1.0	2.5	1.1	2.9	1.0	3.0	1.0	2.8	0.9	2.8	0.9
10. Summarize multiple evaluations of a student during a course into a single digit or letter	2.9	1.0	2.4	0.9	2.8	1.0	3.1	1.0	3.0	0.9	3.0	0.9

Table 3 (continued)

Uses of Grades	Faculty Education (N=75) $\bar{X}$	SD	Faculty Non-Education (N=29) $\bar{X}$	SD	Faculty Total (N=104) $\bar{X}$	SD	Students Education (N=113) $\bar{X}$	SD	Students Non-Education (N=336) $\bar{X}$	SD	Students Total (N=449) $\bar{X}$	SD
11. Provide students with feedback on their course performance	2.1	1.6	1.8	0.9	2.0	1.4	2.4	1.1	2.1	1.0	2.2	1.0
12. Provide a method for maintaining academic standards	2.7	1.2	2.1	1.1	2.6	1.2	2.5	1.0	2.4	1.0	2.4	1.0

[1]1=very important, 2=important, 3=somewhat important, 4=not important.

analysis of the means or carried out a profile analysis. For our purposes, a less sophisticated analysis seemed sufficient. For faculty, a difference of about .4 was needed between two means for the difference to be statistically significant at the .05 level of confidence. For the students, because of the larger sample size, the required difference was about .2. These differences were used as guidelines for discussing the observed mean differences.

Our first observation was the high level of agreement between faculty and student ratings. The rank order correlation between the two sets of mean ratings was about .74. However, students did tend to rate higher than faculty the importance of the 12 uses of grades (for 8 of 12 comparisons, the students rated the use of grades as more important than the faculty did). A difference between faculty and students of as much as ½ point in the mean ratings was obtained for only three of the uses (uses 4, 5, 6). With respect to these three uses, students more than faculty tended to feel that it was important for grades to provide information about a faculty member's teaching methods, curriculum, and competence. (As an aside we noted that there were differences between School of Education and non-school of Education personnel in the overall ratings of the importance of the 12 uses of grading. Generally, School of Education personnel [faculty and students] tended to rate the various uses of grades as less important than their counterparts.)

Which uses of grades were seen as the most important? For the faculty, the answer was clear (listed in order of importance):

11. Provide students with feedback on their course performance.
8. Inform others (e.g., advisors, future teachers, employers, registrar) about student performance.
1. Motivate students to do good work in the course.
2. Provide instructors with information about student progress.
3. Improve a student's ability to critically assess his or her own work.

These five uses of grading were rated as significantly more important than the other uses of grades on our list (with the exception of use 12--maintaining academic standards). For the students, the most important uses of grades were less clear because of the closeness of all of the mean ratings. Basically though, faculty and students agreed on the most important uses of grades.

Acceptability of Five Grading Systems for Accomplishing 12 Possible Uses of Grades

In Table 4 are reported the percentages of "acceptable ratings" given by faculty and students for the five grading systems with respect to the 12 possible uses of grading. Respondents used a four point rating scale:

1=highly acceptable, 2=acceptable, 3=minimally acceptable, 4=unacceptable

An "acceptable rating," for the purposes of this analysis was defined as a "1," "2," or "3" rating.

Consistent with the earlier reported results, students tended to rate the grading systems (regardless of the use) as more acceptable than the faculty. Across the 60 pairs of percentages that could be compared (12 uses of grades x 5 grading systems), the student percentage (reflecting acceptability) was higher, 55 times. All five exceptions to the pattern occurred with criterion-referenced grading. Here, faculty rated the merits of criterion-referenced grading higher (albeit, only slightly higher) than the students.

The most important comparisons concern the relative merits of the five grading systems. For 11 of the 12 uses of grades, more faculty found criterion-referenced grading to be acceptable than any other. For students, criterion-referenced grading received the highest ratings, 10 of the 12 times.

Next, we did an indepth analysis of the response data for the five uses of grades (1, 2, 3, 8, 11) identified in the last section as the

Table 4

Faculty and Student Ratings of the Acceptability
of Five Grading Systems for Accomplishing
12 Possible Uses of Grades
(Percentage of "Acceptable" Ratings Are Reported)

Uses of Grades	Grading System	Faculty			Students		
		Education (N=75)	Non-Education (N=29)	Total (N=104)	Education (N=113)	Non-Education (N=336)	Total (N=449)
1. Motivate students to do good work in the course	P/F	63.1	63.6	63.2	75.7	74.6	74.9
	P/NR	46.2	26.1	40.9	50.5	51.2	50.9
	H/P/F	90.8	91.7	91.0	95.0	93.4	93.8
	NRG	68.7	87.5	73.6	78.8	86.8	85.6
	CRG	83.6	100.0	88.0	84.2	90.7	89.1
2. Provide instructors with information about student progress	P/F	48.5	38.1	46.0	67.0	60.1	61.9
	P/NR	39.4	22.7	35.2	56.6	52.7	53.5
	H/P/F	86.4	60.9	79.8	93.9	90.1	91.0
	NRG	73.1	87.0	76.7	82.8	89.2	87.7
	CRG	91.0	100.0	93.4	89.1	93.0	92.1
3. Improve a student's ability to critically assess his or her own work	P/F	46.9	36.4	44.2	66.3	60.3	61.5
	P/NR	42.2	26.1	37.9	50.0	52.5	51.8
	H/P/F	87.5	62.5	70.7	88.0	86.3	86.7
	NRG	72.7	83.3	75.6	74.7	88.2	85.0
	CRG	92.4	96.0	93.4	87.3	89.9	89.3
4. Provide instructors with ininformation about their teaching methods	P/F	38.8	23.8	35.2	52.0	58.5	58.7
	P/NR	38.8	27.3	36.0	51.0	50.6	50.6
	H/P/F	70.1	43.5	63.3	80.0	84.7	83.6
	NRG	61.2	69.6	63.3	76.5	88.5	85.6
	CRG	79.1	79.2	79.1	83.8	88.9	87.7

Table 4 (cont)

Uses of Grades	Grading System	Faculty			Students		
		Education (N=75)	Non-Education (N=29)	Total (N=104)	Education (N=113)	Non-Education (N=336)	Total (N=449)
5. Provide instructors with information about their curriculum	P/F	47.6	19.0	39.5	60.6	61.8	61.6
	P/NR	39.7	14.3	33.3	53.5	54.2	54.1
	H/P/F	57.1	36.4	51.8	75.8	81.3	80.0
	NRG	55.4	63.6	57.5	77.8	84.9	83.2
	CRG	70.8	69.6	70.5	81.8	85.1	84.3
6. Assess teacher competence	P/F	41.9	27.3	38.1	54.1	55.7	55.4
	P/NR	35.5	27.3	33.3	46.4	51.6	50.5
	H/P/F	48.4	43.5	47.1	68.4	78.6	76.2
	NRG	46.9	52.2	48.3	71.1	83.0	80.2
	CRG	71.9	75.0	72.7	81.6	84.2	83.7
7. Rank students on their performance in the course	P/F	31.2	31.8	31.4	55.0	54.8	54.7
	P/NR	28.1	21.7	26.4	45.5	46.3	46.0
	H/P/F	78.1	58.3	72.7	79.2	87.9	85.9
	NRG	81.8	83.3	82.8	83.0	92.0	89.9
	CRG	80.0	96.0	84.6	82.0	91.2	89.0
8. Inform others (e.g., advisors, future teachers, employers, registrar) about student performance	P/F	53.0	40.9	50.0	54.0	54.3	54.3
	P/NR	42.8	26.1	38.2	43.0	40.4	40.9
	H/P/F	92.4	70.8	86.7	84.0	87.9	87.0
	NRG	76.8	79.2	77.4	78.8	92.4	89.1
	CRG	88.1	96.0	90.2	86.3	90.9	89.8
9. Compare student performance against absolute standards of performance	P/F	40.3	18.2	34.5	48.5	47.1	47.5
	P/NR	32.4	13.0	27.1	42.9	44.6	44.0
	H/P/F	75.8	45.8	68.6	73.7	77.1	76.3
	NRG	75.4	62.5	60.7	79.8	82.6	82.0
	CRG	79.4	92.0	86.7	82.8	87.5	86.4

Table 4 (cont)

Uses of Grades	Grading System	Faculty			Students		
		Education (N=75)	Non-Education (N=29)	Total (N=104)	Education (N=113)	Non-Education (N=336)	Total (N=449)
10. Summarize multiple evaluations of a student during a course into a single digit or letter	P/F	47.0	40.9	45.5	52.0	51.8	52.0
	P/NR	42.4	34.8	40.4	47.5	45.7	46.2
	H/P/F	75.8	54.2	70.0	73.7	72.3	72.7
	NRG	75.4	87.5	78.5	81.8	85.5	84.6
	CRG	79.4	96.0	83.9	85.9	89.1	88.4
11. Provide students with feedback on their course performance	P/F	55.9	45.5	53.7	65.0	62.9	63.5
	P/NR	50.0	34.8	46.2	62.6	56.2	57.9
	H/P/F	88.2	75.0	84.8	86.9	91.0	90.0
	NRG	75.4	91.7	79.6	83.7	92.3	90.3
	CRG	89.7	100.0	92.5	90.0	92.1	91.6
12. Provide a method for maintaining academic standards	P/F	40.9	31.8	38.6	62.0	56.7	57.8
	P/NR	34.8	8.7	28.1	56.6	46.0	48.4
	H/P/F	72.7	66.7	71.1	86.0	85.1	85.3
	NRG	63.6	69.2	67.8	82.7	88.5	87.1
	CRG	78.8	92.0	82.4	86.9	90.2	89.4

most important. The chart below summarizes the average percentage (across the five most important uses of grades) of "acceptable" ratings given to the five grading systems by faculty and students:

Grading System	Faculty	Students
P/F	51.4	63.2
P/NR	39.7	51.0
H/P/F	82.6	89.7
NRG	76.6	87.5
CRG	91.5	90.4

The chart above reveals several things. One, for the five most important uses of grades as identified by the faculty and students, P/F grading and P/NR grading are clearly unacceptable alternatives. Two, there is not a great deal to choose among the other three as far as the students were concerned. Over 87% of the students found all three grading systems at least minimally acceptable. Three, for faculty, CRG was preferred to the other two most popular (NRG and H/P/F) by a statistically significant margin ($p<.05$). Perhaps the most surprising result revealed by the chart was the "acceptability" of norm-referenced grading to so many students.

Effects of Different Systems of Grading on Various Course Outcomes

Reported in Table 5 are the percentages of faculty and students rating positive and negative effects of five grading systems on a variety of course outcomings. In Table 6 are results that bear on the question of the course outcomes that faculty and students consider to be most important.

Which grading systems were rated by faculty and students as most likely "to help" or "to hinder" the accomplishment of the course outcomes?

Table 5

Effect of Different Systems of Grading
on Various Course Outcomes

Course Outcomes	Grading System	Effect ("+" or "-")	Faculty			Students		
			Education (N=75)	Non-Education (N=29)	Total (N=104)	Education (N=113)	Non-Education (N=336)	Total (N=449)
1. Minimize student competition	P/F	+	56.8	58.6	57.3	70.8	63.9	65.5
		-	1.4	0.0	1.0	3.5	5.1	4.7
	H/P/NR	+	16.2	10.3	14.6	37.2	39.0	38.5
		-	10.8	6.9	9.7	19.5	12.9	14.5
	P/NR	+	71.6	48.3	65.0	72.6	71.9	71.9
		-	1.4	3.4	1.9	4.4	5.7	5.4
	NRG	+	2.7	0.0	1.9	8.0	8.1	8.1
		-	77.0	55.2	70.9	73.2	71.3	71.8
	CRG	+	20.3	3.4	15.5	23.9	23.7	23.9
		-	20.3	44.8	27.2	39.8	42.5	41.7
2. Maximize student enthusiasm	P/F	+	13.3	7.1	11.7	31.9	16.6	20.4
		-	16.0	21.4	17.5	26.5	37.3	34.5
	H/P/NR	+	38.7	21.4	34.0	52.2	45.0	46.8
		-	8.0	7.1	7.8	11.5	14.1	13.4
	P/NR	+	21.3	7.1	17.5	25.7	20.7	21.9
		-	20.0	35.7	24.3	38.9	47.1	45.2
	NRG	+	4.0	21.4	8.7	28.3	37.8	35.3
		-	46.7	21.4	39.8	38.9	24.9	28.4
	CRG	+	30.7	39.3	33.0	46.9	45.3	45.9
		-	14.7	10.7	13.6	17.7	22.5	21.3

Table 5 (contd)

Course Outcomes	Grading System	Effect ("+" or "-")	Faculty			Students		
			Education (N=75)	Non-Education (N=29)	Total (N=104)	Education (N=113)	Non-Education (N=336)	Total (N=449)
3. Maximize student performance	P/F	+	14.7	3.4	11.5	23.0	9.0	12.5
		-	18.7	41.4	25.0	32.7	45.5	42.2
	H/P/NR	+	30.7	13.8	26.0	47.3	35.6	38.5
		-	2.7	6.9	3.8	13.4	14.7	14.3
	P/NR	+	14.7	6.9	12.5	19.5	14.1	15.4
		-	36.0	51.7	40.4	50.4	61.4	58.7
	NRG	+	24.0	31.0	26.0	38.9	47.6	45.3
		-	24.0	10.3	20.3	27.4	16.2	19.0
	CRG	+	52.0	69.0	56.7	59.3	61.4	60.9
		-	6.7	6.9	6.7	13.3	12.0	12.3
4. Minimize cheating	P/F	+	42.7	35.7	40.8	60.2	53.6	55.1
		-	1.3	3.6	1.9	6.2	10.5	9.4
	H/P/NR	+	17.3	21.4	18.4	34.5	38.3	37.3
		-	4.0	3.6	3.9	12.4	12.3	12.3
	P/NR	+	74.7	53.6	68.9	66.4	67.4	67.0
		-	2.7	0.0	1.9	7.1	9.0	8.5
	NRG	+	2.7	0.0	1.9	9.7	13.8	12.7
		-	70.7	39.3	62.1	63.7	55.7	57.8
	CRG	+	10.7	14.3	11.7	19.5	17.7	18.3
		-	29.3	32.1	30.1	44.2	50.6	48.9

Table 5 (contd)

Course Outcomes	Grading System	Effect ("+" or "-")	Faculty			Students		
			Education (N=75)	Non-Education (N=29)	Total (N=104)	Education (N=113)	Non-Education (N=336)	Total (N=449)
5. Maximize outside reading	P/F	+	10.7	3.4	8.7	22.1	14.7	16.5
		-	21.3	37.9	26.0	29.2	41.6	38.4
	H/P/NR	+	38.7	20.7	33.7	40.7	38.9	39.3
		-	5.3	6.9	5.8	14.2	15.9	15.4
	P/NR	+	10.7	6.9	9.6	19.5	15.9	16.8
		-	48.0	48.3	48.1	50.4	58.0	56.2
	NRG	+	25.3	17.2	23.1	29.2	42.8	39.3
		-	8.0	13.8	9.6	23.9	18.6	19.9
	CRG	+	50.7	72.4	56.7	59.3	55.1	56.3
		-	6.7	3.4	5.8	15.0	14.7	14.7
6. Develop positive self-image	P/F	+	17.3	6.9	14.4	37.2	21.3	25.2
		-	12.0	24.1	15.4	15.9	32.3	28.1
	H/P/NR	+	37.3	20.7	32.7	52.2	47.9	48.9
		-	4.0	6.9	4.8	6.2	9.9	8.9
	P/NR	+	26.7	13.8	23.1	35.4	22.8	25.9
		-	20.0	27.6	22.1	25.7	37.7	34.8
	NRG	+	5.3	13.8	7.7	20.4	35.0	31.3
		-	27.3	20.7	32.7	50.4	31.4	36.2
	CRG	+	37.3	51.7	41.3	45.1	50.6	49.3
		-	12.0	10.3	11.5	23.0	21.0	21.4

Table 5 (contd)

Course Outcomes	Grading System	Effect ("+" or "-")	Faculty			Students		
			Education (N=75)	Non-Education (N=29)	Total (N=104)	Education (N=113)	Non-Education (N=336)	Total (N=449)
7. Minimize course dropouts	P/F	+	41.3	31.0	38.5	51.3	54.8	53.8
		-	0.0	0.0	0.0	7.1	5.4	5.8
	H/P/NR	+	25.3	10.3	21.2	38.9	45.2	43.5
		-	0.0	3.4	1.0	8.8	6.0	6.7
	P/NR	+	66.7	55.2	63.5	67.3	75.4	73.2
		-	2.7	3.4	2.9	10.6	5.7	6.9
	NRG	+	2.7	6.9	3.8	10.6	10.5	10.5
		-	50.7	34.5	46.2	54.9	52.7	53.3
	CRG	+	14.7	3.4	11.5	21.2	15.0	16.7
		-	28.0	44.8	32.7	36.3	52.4	48.2
8. Provide data for course evaluation	P/F	+	9.3	3.6	7.8	21.2	13.8	15.7
		-	24.0	32.1	26.2	27.4	42.3	38.5
	H/P/NR	+	16.0	7.1	13.6	32.7	32.2	32.3
		-	8.0	7.1	7.8	14.2	16.6	15.9
	P/NR	+	13.3	7.1	11.7	18.6	13.0	14.3
		-	46.7	42.9	45.6	51.3	59.0	57.2
	NRG	+	20.0	25.0	21.4	40.7	56.3	52.2
		-	12.0	7.1	10.7	14.2	12.0	12.6
	CRG	+	65.3	64.3	65.0	62.8	60.2	61.0
		-	1.3	0.0	1.0	10.6	9.9	10.1

Table 5 (contd)

Course Outcomes	Grading System	Effect ("+" or "-")	Faculty			Students		
			Education (N=75)	Non-Education (N=29)	Total (N=104)	Education (N=113)	Non-Education (N=336)	Total (N=449)
9. Maintain academic standards	P/F	+	6.7	0.0	4.8	19.5	11.7	13.6
		-	26.7	41.4	30.8	29.2	43.1	39.5
	H/P/NR	+	20.0	3.4	15.4	35.4	37.1	36.6
		-	5.3	13.8	7.7	15.9	16.2	16.1
	P/NR	+	12.0	3.4	9.6	17.7	11.1	12.7
		-	52.0	48.3	51.0	54.9	66.8	63.8
	NRG	+	22.7	27.6	24.0	44.2	50.3	48.7
		-	13.3	10.3	12.5	15.9	12.9	13.6
	CRG	+	54.7	62.1	56.7	60.2	66.2	64.7
		-	4.0	3.4	3.8	12.4	8.7	9.6
10. Maximize class attendance	P/F	+	5.3	10.3	6.7	20.4	11.1	13.4
		-	18.7	24.1	20.2	33.6	46.7	43.3
	H/P/NR	+	20.0	10.3	17.3	27.4	31.7	30.6
		-	6.7	10.3	7.7	23.9	18.3	19.6
	P/NR	+	5.3	10.3	6.7	15.0	11.1	12.1
		-	52.0	48.3	51.0	59.3	66.8	65.0
	NRG	+	32.0	17.2	27.9	48.7	56.0	54.0
		-	9.3	17.2	11.5	18.6	8.1	10.7
	CRG	+	52.0	65.5	55.8	54.9	64.7	62.3
		-	5.3	10.2	6.7	12.4	9.9	10.5

Table 5 (contd)

Course Outcomes	Grading System	Effect ("+" or "-")	Faculty			Students		
			Education (N=75)	Non-Education (N=29)	Total (N=104)	Education (N=113)	Non-Education (N=336)	Total (N=449)
11. Maximize expressions of personal opinions	P/F	+	31.1	25.0	29.4	45.1	35.2	37.7
		-	4.1	10.7	5.9	14.2	21.1	19.3
	H/P/NR	+	35.1	25.0	32.4	41.6	45.2	44.2
		-	6.8	7.1	6.9	12.4	8.4	9.4
	P/NR	+	45.9	28.6	41.2	45.1	39.5	40.8
		-	16.2	10.7	14.7	21.2	28.6	26.9
	NRG	+	9.5	14.3	10.8	19.5	32.5	29.1
		-	35.1	28.6	33.3	46.9	33.4	36.8
	CRG	+	28.4	28.6	28.4	31.9	40.4	38.3
		-	16.2	25.0	18.6	30.1	30.1	30.0
12. Maximize amount of learning	P/F	+	14.9	3.4	11.7	27.4	14.7	17.9
		-	16.2	31.0	20.4	26.5	36.8	34.2
	H/P/NR	+	28.4	10.3	23.3	44.2	38.6	40.0
		-	5.4	10.3	6.8	10.6	16.8	15.2
	P/NR	+	18.9	6.9	15.5	22.1	14.7	16.5
		-	39.2	55.2	43.7	46.0	57.5	54.7
	NRG	+	16.2	24.1	18.4	26.5	43.4	39.1
		-	23.0	13.8	20.4	30.1	18.6	21.4
	CRG	+	52.7	65.5	56.3	49.6	58.7	56.5
		-	6.8	6.9	6.8	20.4	13.2	15.0

Table 5 (contd)

Course Outcomes	Grading System	Effect ("+" or "-")	Faculty			Students		
			Education (N=75)	Non-Education (N=29)	Total (N=104)	Education (N=113)	Non-Education (N=336)	Total (N=449)
13. Maximize opportunity for further study	P/F	+	12.2	6.9	10.7	35.4	21.9	25.3
		-	8.1	27.6	13.6	21.2	31.5	28.9
	H/P/NR	+	44.6	27.6	39.8	49.6	42.9	44.5
		-	2.7	3.4	2.9	11.5	11.7	11.6
	P/NR	+	21.6	10.3	18.4	24.8	25.8	25.5
		-	25.7	37.9	29.1	41.6	44.1	43.6
	NRG	+	13.5	10.3	12.6	25.7	39.3	35.8
		-	27.0	17.2	24.3	30.1	21.9	23.9
	CRG	+	33.8	48.3	37.9	45.1	52.9	51.0
		-	9.5	13.8	10.7	18.6	16.5	17.0

Table 6

Percentage of Faculty and Students Indicating the Importance of Various Course Outcomes

Course Outcomes	Faculty			Students		
	Education (N=53)	Non-education (N=24)	Total (N=77)	Education (N=51)	Non-education (N=258)	Total (N=309)
1. Minimize student competition	20	25	22	34	27	28
2. Maximize student enthusiasm	95	82	93	98	77	80
3. Maximize student performance	87	94	89	94	79	82
4. Minimize cheating	15	30	20	29	20	22
5. Maximize outside reading	67	34	56	71	51	54
6. Develop positive self-image	73	48	67	91	72	75
7. Minimize course dropouts	8	5	7	17	9	11
8. Provide data for course evaluation	24	21	23	13	21	20
9. Maintain academic standards	49	30	43	47	40	41
10. Maximize class attendance	30	33	31	41	24	27
11. Maximize expressions of personal opinions	70	52	66	79	58	61
12. Maximize amount of learning	98	90	96	100	97	98
13. Maximize opportunity for further study	78	59	72	91	67	71

The most significant results are summarized below:

Course Outcome	—Grading Systems Which— Help	Hinder
1. Minimize student competition	P/F, P/NR	NRG
2. Maximize student enthusiasm	H/P/F, CRG	P/NR
3. Maximize student performance	CRG	P/NR
4. Minimize cheating	P/F, P/NR	NRG, CRG
5. Maximize outside reading	H/P/F, CRG	P/NR
6. Develop positive self-image	H/P/F, CRG	
7. Minimize course dropouts	P/F, P/NR	NRG, CRG
8. Provide data for course evaluation	CRG	P/NR
9. Maintain academic standards	CRG	P/NR
10. Maximize class attendance	CRG, NRG	P/NR
11. Maximize expressions of personal opinions	H/P/F, P/NR	
12. Maximize amount of learning	CRG	P/NR
13. Maximize opportunity for further study	H/P/F, CRG	

Of course, there were some differences in the ratings of faculty and students, but the above summary essentially "captures" the feelings of both groups.

Next, we conducted an analysis of the response data for the three most important course outcomes (2, 3, 12) revealed by Table 6. These are:

2. Maximize student enthusiasm
3. Maximize student performance
12. Maximize amount of learning.

The chart below summarizes the average percentage (across the three most important course outcomes) of "+" and "-" effects ascribed to the five grading systems by faculty and students:

Grading System	Effect	Faculty	Students
P/F	+	11.6	16.9
H/P/F	+	27.7	41.7
P/NR	+	15.1	17.9
NRG	+	17.7	29.8
CRG	+	48.6	54.4
P/F	-	20.9	36.9
H/P/F	-	6.1	14.3
P/NR	-	36.1	52.8
NRG	-	26.8	22.9
CRG	-	9.0	16.2

The chart above leads to this conclusion: Across the three most important course outcomes as identified by faculty and students, criterion-referenced grading and honors, pass, fail grading were seen as the most desirable and pass/fail grading and pass/no record were seen as the least desirable.

Conclusion

The controversy over grading purposes and practices in Higher Education will not subside until institutions approach the study of the problem in a more systematic way. While we do not intend to offer a comprehensive plan for studying the problem of grading purposes and practices in this paper (see Hambleton & Rovinelli, in press), several of the activities outlined in this paper would be part of any comprehensive plan. The views of faculty and students toward a variety of purposes or uses of grades in different instructional settings should be considered. Second, the relative merits of possible grading systems for accomplishing the most important of

the uses of grades in a particular institution should be studied. How faculty and student data will be used in any institution will depend on the types of flexibility in grading practices that are possible. Finally, the relative effects of different grading systems on various course outcomes should be considered. Of course, the above steps are three of many that need to be considered in developing a college grading policy.

While the results from this study cannot be generalized to other institutions, we feel that the study provides some useful guidelines for conducting grading research studies. Three major shortcomings of many previous studies on grading were overcome. First, there was recognition of the fact that there are many uses of grades, and the importance of each use of grades will sometimes depend on the particular instructional setting. It is inappropriate to ask individuals about what they believe to be the most important uses of grades without first clearly specifying the instructional setting. In our study, instructional setting was not a significant factor in faculty and student ratings, but it may be elsewhere. Second, a study of the merits of different grading systems should be conducted relative to the various possible uses of grades. Finally, a comprehensive review of grading purposes and practices should include <u>both</u> faculty and student views.

The results of this study revealed several things. First, "instructional setting" had no effect on the ratings of the importance of 12 common uses of grades. Second, five uses of grades were rated as significantly more important than others we studied. They were:

Inform others (e.g., advisors, future teachers, employers, registrar) about student performance

Provide students with feedback on their course performance

Motivate students to do good work in the course

Provide instructors with information about student progress

Improve a student's ability to critically assess his or her own work.

Three, three course outcomes were viewed by faculty and students as more important than others. These were:

Maximize student enthusiasm
Maximize student performance
Maximize amount of learning.

If an institution were to select a single grading system, it ought to choose one that would be acceptable to both faculty and students for accomplishing the most important uses of grades. Also, the grading system should have a positive effect on the most important course outcomes. From our data, we found that over 90% of faculty and students found criterion-referenced grading to be "at least minimally acceptable." No other grading system was rated as highly (H/P/F was second best). With respect to the most important course outcomes, again, criterion-referenced grading was rated to be best. Pass/fail grading was considered to be the least acceptable. (This is especially important to us because the School of Education is currently using a pass/fail grading system!)

On the basis of our results, a recommendation to adopt a criterion-referenced grading system would seem reasonable. Still, it is easier said than done, for of all the grading systems, criterion-referenced grading may be the most difficult system to implement properly. For one, criterion-referenced grading requires instructors to specify their course outcomes in rather specific terms. Second, test development methods require careful attention by instructors to be sure that test items measure course objectives and that the examinations have content validity. Finally, there is the problem of setting "performance standards" to separate "A" level performance, from "B" level performance and so on.

References

Baucom, T. V. Evaluation of college students. Improving College and University Teaching, 1974, 22, 27.

Ericksen, S. C., & Bluestone, B. Z. Grading ≠ Evaluation. Memo to the faculty. Technical Report No. 46. Ann Arbor: University of Michigan, 1971.

Hales, L. W., Bain, P. T., & Rand, O. P. The Pass-fail option: The congruence between the rationale for the student reasons in electing. Journal of Educational Research, 1973, 66, 295-298.

Hambleton, R. K., & Novick, M. R. Toward an integration of theory and method for criterion-referenced tests. Journal of Educational Measurement, 1973, 10, 159-170.

Hambleton, R. K., & Rovinelli, R. J. Toward better college grading practices: A framework for research and development. Improving College and University Teaching, in press.

Hambleton, R. K., & Rovinelli, R. J. Toward better achievement tests and test score interpretations in PSI courses. Laboratory of Psychometric and Evaluative Research Report No. 25. Amherst, Mass.: School of Education, University of Massachusetts, 1976.

Hunt, R. A. Student grades as a feedback system: The case for a confidential multiple grade. Measurement and Evaluation in Guidance, 1972, 5, 345-359.

Karlins, M., Kaplan, M., & Stuart, W. Academic attitudes and performance as a function of different grading systems: An evaluation of Princeton's pass-fail grading system. Journal of Experimental Education, 1969, 37, 38-50.

Millman, J. Criterion-referenced measurement. In W. J. Popham (Ed.), Evaluation in education: Current practices. San Francisco: McCutchan Publishers, 1974.

Murray, J. Faculty and student views of college grading purposes and practices (Tent. title). Unpublished doctoral dissertation, University of Massachusetts, Amherst, 1977.

Reiner, J. R., & Jung, L. B. Enrollment patterns and academic performance as a function of registration under a pass-fail grading system. Interchange, 1972, 3, 53-62.

Scriven, M. The evaluation of students. Unpublished manuscript, 1976.

Stallings, W. M., & Smock, H. R. The pass-fail grading option at a state university: A five semester evaluation. Journal of Educational Measurement, 1971, 8, 153-160.

Stallings, W. M., Wolff, J. L., & Maehr, M. L. Fear of failure and the pass-fail grading option. Journal of Experimental Education, 1969, 38, 87-91.

Stevens, E. I. Grading systems and student mobility. Educational Record, 1973, 54, 308-312.

Warren, J. R. College grading practices: An overview. Research Bulletin No. 71-12. Princeton, New Jersey: Educational Testing Service, 1971.

Williams, R. G., & Miller, H. G. Grading students: A failure to communicate. Clearing House, 1973, 47, 332-337.